IDENTIFICATION OF BIOLOGICAL (MICRO)ORGANISMS BY DETECTION OF THEIR HOMOLOGOUS NUCLEOTIDE SEQUENCES ON ARRAYS

Field of the invention

The present invention is in the field of diagnosis and is related to a method and kit comprising reagents and means for the identification (detection and/or quantification) of (micro)organisms among other ones having homologous nucleotide sequences by identification of their nucleotide sequences, after amplification by a single primer pair.

The invention is especially suited for the identification and/or quantification of (micro)organisms of the same genus or family or for the detection and/or quantification of related genes in a specific (micro)organism present in a biological sample.

Background of the invention

The development of the biochips technology allows the detection of multiple nucleotide sequences simultaneously in a given assay and thus allow the identification of the corresponding organism or part of the organism. Arrays are solid supports containing on their surface a series of discrete regions bearing capture nucleotide sequences (or probes) that are able to bind (by hybridisation) to a corresponding target nucleotide sequence(s) possibly present in a sample to be analysed. If the target sequence is labelled with modified nucleotides during a reverse transcription or an amplification of said sequence, then a signal can be detected and measured at the binding location. Its intensity gives an estimation of the amount of target sequences present in the sample. Such technology allows the identification and/or quantification of genes or species for diagnostic or screening purpose.

State of the art

The Company Affymetrix Inc. has developed a method for direct synthesis of oligonucleotides upon a solid support, at specific locations by using masks at each step of the processing. Said method comprises the addition of a new nucleotide on a growing oligonucleotide in order to obtain a desired sequence at a desired location. This method is derived from the photolithographic technology and is coupled with the use of photoprotective groups, which are released before a new nucleotide is added (EP-A1-0476014, US-A-5,445,934, US-A-5,143,854 and US-5,510,270). However, only small oligonucleotides are present on the surface, and said method finds applications mainly for sequencing or identifying a pattern of positive spots corresponding to each specific oligonucleotide bound on the array. The characterization of a target sequence is obtained by comparison of such pattern with a reference. Said technique was applied to the identification of Mycobacterium tuberculosis rpoB gene (WO97/29212 and WO98/28444), wherein the capture nucleotide sequence comprises less than 30 nucleotides and from the analysis of two different sequences that may differ by a single nucleotide (the identification of SNPs or genotyping). Small capture nucleotide sequences (having a length comprised between 10 and 20 nucleotides) are preferred since the discrimination between two oligonucleotides differing in one base is higher, when their length is smaller.

The lack of sensitivity of the method is illustrated by the fact that it cannot detect directly amplicons resulting from genetic amplification (PCR). A double amplification with primer(s) bearing a T3 or T7 sequences and then a retrotranscription with a RNA polymerase. These RNA are cut into pieces of about 40 bases before being detected on an array (example 1 of WO 97/29212). However, long DNA or RNA fragments hybridize very slowly on

capture probes present on a surface. Said methods are therefore not suited for the detection of homologous sequences since the homology varies along the sequences and so part of the pieces could hybridize on the same capture probes. Therefore, a software for the interpretation of the results should be incorporated in the method for allowing interpretation of the obtained data.

However, for gene expression array which is based on the cDNA copy of mRNA the same problem is encountered when using small capture probe arrays: the rate of hybridisation is low. Therefore, the fragments are cut into smaller species and the method requires the use of several capture nucleotide sequences in order to obtain a pattern of signals which attest the presence of a given gene (WO97/10364 and WO97/27317). Said cutting also decreases the number of labelled nucleotides, and thus reduces the obtained signal. In this case, the use of long capture nucleotide sequences give a much better sensitivity to the detection. In the many gene expression applications, the use of long capture probes is not a problem, when cDNA to be detected originates from genes having different sequences, since there is no cross-reactions between them. Long capture nucleotide sequences give the required sensitivity, however, they will hybridize to other homologous sequences.

Using membranes or nylon supports are proposed to increase the sensitivity of the detection on solid support by incorporation of a spacer between the support and the capture nucleotide sequences. Van Ness et al. (Nucleic Acids Research, Vol.19, p.3345, 1991) describe a poly(ethyleneimine) arm for the binding of DNA on nylon membranes. The European patent application EP-0511559 describes a hexaethylene glycol dervivative as spacer for the binding of small oligonucleotides upon a membrane. When membranes like nylon are used as support, there is no control of the site of binding between the solid support and the

oligonucleotides and it was observed that a poly dT tail increased the fixation yield and so the resulting hybridization (W089/11548). Similar results are obtained with repeated capture sequences present in a polymer (US 5,683,872).

Guo et al. (Nucleic Acids Research 22, 5456, 1994) teach the use of poly dT of 15 bases as spacer for the binding of oligonucleotides on glass with increased sensitivity of hybridization.

The document WO99/16780 describes the detection of 4 homologous sequences of the gene femA on nylon strips. However, no data on the sensitivity of the method and the detection is presented. In said document, the capture nucleotide sequences comprise between 15 and 350 bases with homology less than 50% with a consensus sequence.

The publication of Anthony et al. (Journal of clinical microbiology, Vol.38 nr.2, p.7817-8820) describes the use of a membrane array for the discrimination with low sensitivity of homologous sequences originated from a several related organisms. Targets to detect are rDNA amplified from bacteria by consensus PCR and the detection is obtained on nylon array containing capture nucleotide sequences for said bacteria and having the capture nucleotide sequences having between 20 and 30 bases which are covalently linked to the nylon, and there is no control of the portion of the sequence which is available for hybridization.

Aims of the invention

The present invention aims to provide a new method and device to improve microarrays or biochips technology for the easy identification (detection and/or quantification) of a large number of (micro)organisms or

portions of (micro)organisms having homologous nucleotide sequences.

A further aim of the invention is to provide such method and device which are based upon a simplified technology requiring the use single primer(s) in an amplification step and which allow the identification (detection and/or quantification) of a specific target sequence by the identification and/or recording of a single spot signal upon said microarray, said signal resulting only from the specific binding of the target sequence with its corresponding capture sequence.

Definitions

The terms "nucleic acid, oligonucleotide, array, probe, target nucleic acid, bind substantially, hybridising specifically to, background, quantifying" are the ones described in the international patent application WO97/27317 incorporated herein by reference.

The terms "nucleotide triphosphate, nucleotide, primer sequence" are those described in the document WO00/72018 and PCT/BE00/00123 incorporated herein by references.

The terms "Homologous sequences" and "consensus sequence" are described in the European patent application 00870055.1 incorporated herein by reference.

Summary of the invention

The inventors have discovered that it is possible to drastically simplify the identification of one or several (micro)organisms among many other ones having homologous sequences by combining a single amplification using common primer pair and an identification of the possible (micro)organism(s) by detecting and possibly recording upon an array the presence of a single signal resulting only from a binding between a capture sequence and

its corresponding target sequence and correlating the presence of said detected target sequence to the identification of a genetic sequence specific of said (micro)organism(s). This means that the method and device according to the invention will allow the easy identification/detection of a specific sequence among other homologous sequences and possibly its quantification (characterisation of the number of copies or presence of said organisms in a biological sample) of a target sequence, said target sequence having a nucleotide sequence specific of said (micro)organisms.

Such identification may be obtained directly, after washing of possible contaminants (unbound sequences), by detecting and possibly recording a single spot signal at one specific location, wherein said capture nucleotide sequence was previously bound and said identification is not a result of an analysis of a specific pattern upon the microarray as proposed in the system of the state of the art. Therefore, said method and device do not necessarily need a detailed analysis of said pattern by an image processing and a software analysis.

This invention was made possible by discovering that target sequences can be discriminated from other homologous ones upon an array with high sensitivity by using bound capture nucleotide sequences composed of at least two parts, one being a spacer bound by a single and advantageously predetermined (defined) link to the support (preferably a non porous support) and the other part being a specific nucleotide sequence able to hybridise with the nucleotide target sequence.

Furthermore, said detection is greatly increased, if high concentrations of capture nucleotide sequences are bound to the surface of the solid support.

The present invention is related to the identification of a target sequence obtained from a

biological (micro)organism or a portion thereof, especially a gene possibly present in a biological sample from at least 4 other homologous (micro)organisms or a portion thereof, said other (micro)organisms could be present in the same biological sample and have homologous nucleotide sequences with the target.

Said identification is obtained firstly by a genetic amplification of said nucleotide sequences (target and homologous sequences) by common primer pairs followed (after washing) by a discrimination between the possible different target amplified. Said discrimination is advantageously obtained by hybridization upon the surface of an array containing capture nucleotide sequences at a given location, specific for a target specific for each (micro)organism to be possibly present in the biological sample and by the identification of said specific target through the identification and possibly the recording of a signal resulting from the specific binding of this target upon its corresponding capture sequence at the expected location (single location signal being specific for the target).

According to the invention, the preferred method for genetic amplification is the PCR using two anti-parallel consensus primers which can recognise all said target homologous nucleotide sequences.

The method according to the invention further comprises the step of correlating the signal of detection (possibly recorded) to the presence of :

- specific (micro)organism(s),
- genetic characteristics of a sequence,
- polymorphism of a sequence,
- diagnostic predisposition or evolution (monitoring) of genetic diseases, including cancer of a patient (including the human) from which the biological sample has been obtained.

Therefore, said (micro)organisms could be present in any biological material including genetic material obtained (virus, fungi, bacteria, plant or animal cell, including the human). The biological sample can be also any culture medium wherein microorganisms, xenobiotics or pollutants are present, as well as such extract obtained from a plant or an animal (including a human) organ, tissue, cell or biological fluid (blood, serum, urine, etc).

The method according to the invention can be performed by using a specific identification (diagnostic and/or quantification) kit or device comprising at least an insoluble solid support upon which are bound single stranded capture nucleotide sequences (preferably bound to the surface of the solid support by a direct covalent link or by the intermediate of a spacer) according to an array with a density of at least 4, preferably at least 10, 16, 20, 50, 100, 1000, 4000, 10 000 or more, different single stranded capture nucleotide sequences/cm^2 insoluble solid support surface, said single stranded capture nucleotide sequences having advantageously a length comprised between about 30 and about 600 bases (including the spacer) and containing a sequence of about 10 to about 60 bases, said sequence being specific for the target (which means that said bases of said sequence are able to form a binding with their complementary bases upon the sequence of the target by complementary hybridisation). Preferably, said hybridisation is obtained under stringent conditions (under conditions well-known to the person skilled in the art).

In the method and kit or device according to the invention, the capture nucleotide sequence is a sequence having between 16 and 600 bases, preferably between 30 and 300 bases, more preferably between 40 and 150 bases and the spacer is a chemical chain of at least 6,8 nm long (of at least 4 carbon chains), a nucleotide sequence of more than 30 bases or is nucleotide derivative such as PMA.

The method, kit and device according to the invention are particularly suitable for the identification of a target, being preferably biological (micro)organisms or a part of it, possibly present in a biological sample where at least 4, 12, 15 or even more homologous sequences are present. Because of the high homology, said sequence can be amplified by common primer(s) so that the identification of the target is obtained specifically by the discrimination following its binding with the corresponding capture nucleotide sequence, previously bound at a given location upon the microarray. The sensitivity can be also greater increased if capture nucleotide sequences are spotted to the solid support surface by a robot at high density according to an array. A preferred embodiment of the invention is to use an amount of capture nucleotide sequences spotted on the array resulting in the binding of between about 0.01 to about 5 pmoles of sequence equivalent/cm^2 of solid support surface.

The kit or device according to the invention may also incorporate various media or devices for performing the method according to the invention. Said kit (or device) can also be included in an automatic apparatus such as a high throughput screening apparatus for the detection and/or the quantification of multiple nucleotide sequences present in a biological sample to be analysed. Said kit or apparatus can be adapted for performing all the steps or only several specific steps of the method according to the invention.

In the method, the kit (device) or apparatus according to the invention, the length of the bound capture nucleotide sequences is preferably comprised between about 30 and about 600 bases, preferably between about 40 and about 400 bases and more preferably between about 40 and about 100 bases. Longer nucleotide sequences can be used if they do not lower the binding yield of the target nucleotide sequences usually by adopting hairpin based secondary structure or by interaction with each other.

If the homology between the sequences to be detected is low (between 30 and 60%), parts of the sequence which are specific in each sequence can be used for the design of specific capture nucleotide sequences binding each of the different target sequences. However, it is more difficult to find part of the sequence sufficiently conserved as to design "consensus" sequences which will amplify or copy all desired sequences. If one pair of consensus primers is not enough to amplify all the homologous sequences, then a mixture of two or more primers pairs is added in order to obtain the desired amplifications. The minimum homologous sequences amplified by the same consensus primer is two, nut there is no limitation to said number.

If the sequences show high degree of homology, higher than 60% and even higher than 90%, then the finding of common sequence for consensus primer is easily obtained, but the choice for specific capture nucleotide sequences become more difficult.

In another preferred embodiment of the invention, the capture nucleotide sequences are chemically synthesised oligonucleotides sequences shorter than 100 bases (easily performed on programmed automatic synthesiser). Such sequences can bear a functionalised group for covalent attachment upon the support, at high concentrations.

Longer capture nucleotide sequences are preferably synthesised by PCR amplification (of a sequence incorporated into a plasmid containing the specific part of the capture nucleotide sequence and the non specific part (spacer)).

In a further embodiment of the invention, the specific sequence of the capture nucleotide sequence is separated from the surface of the solid support by at least about 6.8 nm long, equivalent to the distance of at least 20 base pair long nucleotides in double helix form.

In the method, kit (device) or apparatus according to the invention, the portion(s) (or part(ies)) of the capture nucleotide sequences complementary to the target is comprised between about 10 and about 60 bases, preferably between about 15 and about 40 bases and more preferably between about 20 and about 30 bases. These bases are preferably assigned as a continuous sequence located at or near the extremity of the capture nucleotide sequence. This sequence is considered as the specific sequence for the detection. In a preferred form of the invention, the sequence located between the specific capture nucleotide sequence and the support is a non specific sequence.

In another embodiment of the invention, a specific nucleotide sequence comprising between about 10 and about 60 bases, preferably between about 15 and about 40 bases and more preferably between about 20 and about 30 bases is located on a capture nucleotide sequence comprising a sequence between about 30 and about 600 bases.

The method, kit (device) or apparatus according to the invention are suitable for the detection and/or the quantification of a target which is made of DNA or RNA, including sequences which are partially or totally homologous upon their total length.

The method according to the invention can be performed even when a target present between an homology (or sequence identity) greater than 30%, greater than 60% and even greater than 80% and other molecules.

In the method, kit (device) or apparatus according to the invention, the capture nucleotide sequences are advantageously covalently bound (or fixed) upon the insoluble solid support, preferably by one of their extremities as described hereafter.

The method according to the invention gives significant results which allows identification (detection and quantification) with amplicons in solutions at

concentration of lower than about 10 nM, of lower than about 1 nM, preferably of lower than about 0.1 nM and more preferably of lower than about 0.01 nM (= 1 fmole/100 μl).

Another important aspect of this invention is to use very concentrate capture nucleotide sequences on the surface. If too low, the yield of the binding is quickly lower and is undetectable. Concentrations of capture nucleotide sequences between about 600 and about 3,000 nM in the spotting solutions are preferred. However, concentrations as low as about 100 nM still give positive results in favourable cases (when the yield of covalent fixation is high or when the target to be detected is single stranded and present in high concentrations). Such low spotting concentrations would give density of capture nucleotide sequence as low as 20 fmoles per cm^2. On the other side, higher density was only limited in the assays by the concentrations of the capture solutions, but concentrations still higher than 3,000 nM give good results.

The use of these very high concentrations and long probes are two unexpected characteristic features of the invention. The theory of DNA hybridisation proposed that the rate of hybridisation between two DNA complementary sequences in solution is proportional to the square root of the DNA length, the smaller one being the limited factor (Wetmur, J.G. and Davidson, N. 1968, J. Mol. Biol. 3, 584). In order to obtain the required specificity, the specific sequences of the capture nucleotide sequences had to be small compared to the target. Moreover, the targets were obtained after PCR amplification and were double stranded so that they reassociate in solution much faster than to hybridise on small sequences fixed on a solid support where diffusion is low thus reducing even more the rate of reaction. It was unexpected to observe a so large increase in the yield of hybridisation with the same short specific sequence.

The amount of a target which "binds" on the spots is very small compared to the amount of capture nucleotide sequences present. So there is a large excess of capture nucleotide sequence and there was no reason to obtain the binding if even more capture nucleotide sequences.

One may perform the detection on the full length sequence after amplification or copy and when labelling is performed by incorporation of labelled nucleotides, more markers are present on the hybridised target making the assay sensitive.

The method, kit and apparatus according to the invention may comprise the use of other bound capture nucleotide sequences, which may have the same characteristics as the previous ones and may be used to identifying a target from another group of homologous sequences (preferably amplified by common primer(s)).

In the microbiological field, one may use consensus primer(s) specific for each family, or genus, of microorganisms and then identify some or all the species of these various family in an array by using capture nucleotide sequences of the invention. Detection of other sequences can be advantageously performed on the same array (i.e. by allowing an hybridisation with a standard nucleotide sequence used for the quantification, with consensus capture nucleotide sequences for the same or different micro-organisms strains, with a sequence allowing a detection of a possible antibiotic resistance gene by micro-organisms or for positive or negative control of hybridisation). Said other capture nucleotide sequences have (possibly) a specific sequence longer than 10 to 60 bases and a total length as high as 600 bases and are also bound upon the insoluble solid support (preferably in the array made with the other bound capture nucleotide sequences related to the invention). A long capture nucleotide sequence may also be present on the array as consensus capture nucleotide sequence for

hybridisation with all sequences of the microorganisms from the same family or genus, thus giving the information on the presence or not of a microorganism of such family, genus in the biological sample.

The same array can also bear capture nucleotide sequences specific for a bacterial group (Gram positive or Gram negative strains or even all the bacteria).

Another application is the detection of homologous genes from a consensus protein of the same species, such as various cytochromes P450 by specific capture nucleotide sequences with or without the presence of a consensus capture nucleotide sequence for all the cytochromes possibly present in a biological sample. Such detection is performed at the gene level by retrotranscription into cDNA.

The solid support according to the invention can be or can be made with materials selected from the group consisting of gel layers, glasses, electronic devices, silicon or plastic support, polymers, compact discs, metallic supports or a mixture thereof (see EP 0 535 242, US 5,736,257, WO99/35499, US 5,552,270, etc). Advantageously, said solid support is a single glass slide which may comprise additional means (barcodes, markers, etc.) or media for improving the method according to the invention.

The amplification step used in the method according to the invention is advantageously obtained by well known amplification protocols, preferably selected from the group consisting of PCR, RT-PCR, LCR, CPT, NASBA, ICR or Avalanche DNA techniques.

Advantageously, the target to be identified is labelled previously to its hybridisation with the single stranded capture nucleotide sequences. Said labelling (with known techniques from the person skilled in the art) is preferably also obtained upon the amplified sequence previously to the denaturation (if the method includes an amplification step).

Advantageously, the length of the target is selected as being of a limited length preferably between 100 and 200 bases, preferably between 100 and 400 bases and more preferably between 100 and 800 bases. This preferred requirement depends on the possibility to find consensus primers to amplify the required sequences possibly present in the sample. Too long target may reallocate faster and adopt secondary structures which can inhibit the fixation on the capture nucleotide sequences.

Detection of genes is also a preferred application of this invention. The detection of homologous genes is obtained by first retrotranscription of the mRNA and then amplification by consensus primers as described in this invention.

According to a further aspect of the present invention, the method, kit (device) or apparatus according to the invention is advantageously used for the identification of different *Staphylococcus* species or variant, preferably the *S. aureus, the S. epidermidis, the S. saprophyticus, the S. hominis or the S. haemolyticus* for homologous organs present together or separately in the biological sample, said identification being obtained by detecting the genetic variants of the *FemA* gene in said different species, preferably by using a common locations in the *FemA* genetic sequence.

Preferably, the primer(s) and the specific portions of said *FemA* sequence used for obtaining amplified products are the ones described hereafter in example 2. These primers have been selected as consensus primers for the amplification of the *FemA* genes of all of the 16 *Staphylococcus* tested and they probably will amplify the *FemA* from all other possible *Staphylococcus* species.

The detection of the 12 MAGE according to the invention is presented in figure 4 an 5 and in example 4. The

array allows to read the MAGE number by observation of the lines positive for signal bearing the specific capture probes(example 5).

The same application was developed for the Receptors Coupled to the G Proteins(RCGP). These receptors bind all sort of ligands and are responsible for the signal transduction to the cytoplasm and very often to the nucleus by modulating the activity of the transcriptional factors. Consensus primers are formed for the various subtypes of RCGP for dopamine and for serotonine and histamine (figures 6 and 7 and example 4). The same is possible for the histamine and other ligands.

The detection of the various HLA types is also one of the applications of the invention (figure 7 and example. HLA are homologous sequences which differ from one individual to the other. The determination of the HLA type is especially useful in tissue transplantation in order to determine the degree of compatibility between the donor and the recipient. It is also a useful parameter for immunisation. Given the large number of subtypes and the close relation between the homologous sequences it was not always possible to perfectly discriminate one sequence among all the other ones and for some of them there was one or two cross-reactions. In these cases, another capture probe was added on the array which gives a reaction with the sequence to be detected and another cross-reaction, in order to make the identification absolute.

There are several forms of Cytochrome P450 which are also homologous sequences. Figure 9 presents the design of the array to identify several cytochromes P450 after retrotranscription and amplification with consensus primers (example 6).

The detection of polymorphism sequences (which can be considered as homologous even if differing by only one base) can be made also by the method according to the

invention. This is especially useful for the Cytochrome P450 since the presence of certain isoforms modifies the metabolism of some drugs.

Another aspect of the present invention is related to any part of biochips or microarray comprising said above described sequences (especially the specific capture nucleotide sequence described in the examples) as well as a general screening method for the identification of a target sequence specific of said microorganisms of family type descriminated from homologous sequences upon any type of microarrays or biochips by any method.

After hybridisation on the array, the target sequences can be detected by current techniques. Without labelling, preferred methods are the identification of the target by mass spectrometry now adapted to the arrays (US-A-5,821,060) or by intercalating agents followed by fluorescent detection(WO97/27329 or Fodor et al., Nature 364, p. 555 (1993)).

The labelled associated detections are numerous. A review of the different labelling molecules is given in W0 97/27317. They are obtained using either already labelled primer or by incorporation of labelled nucleotides during the copy or amplification step. A labelling can also be obtained by ligating a detectable moiety onto the RNA or DNA to be tested (a labelled oligonucleotide, which is ligated, at the end of the sequence by a ligase). Fragments of RNA or DNA can also incorporate labelled nucleotides at their 5'OH or 3'OH ends using a kinase, a transferase or a similar enzyme.

The most frequently used labels are fluorochromes like Cy3, Cy5 and Cy7 suitable for analysing an array by using commercially available array scanners (General Scanning, Genetic Microsystem,...). Radioactive labelling, cold labelling or indirect labelling with small molecules recognised thereafter by specific ligands (streptavidin or

antibodies) are common methods. The resulting signal of target fixation on the array is either fluorescent, colorimetric, diffusion, electroluminescent, bio- or chemiluminescent, magnetic, electric like impedometric or voltametric (US-A-5,312,527). A preferred method is based upon the use of the gold labelling of the bound target in order to obtain a precipitate or silver staining which is then easily detected and quantified by a scanner.

Quantification has to take into account not only the hybridisation yield and detection scale on the array (which is identical for target and reference sequences) but also the extraction, the amplification (or copying) and the labelling steps.

The method according to the invention may also comprise means for obtaining a quantification of target nucleotide sequences by using a standard nucleotide sequence (external or internal standard) added at known concentration. A capture nucleotide sequence is also present on the array so as to fix the standard in the same conditions as said target (possibly after amplification or copying); the method comprising the step of quantification of a signal resulting from the formation of a double stranded nucleotide sequence formed by complementary base pairing between the capture nucleotide sequences and the standard and the step of a correlation analysis of signal resulting from the formation of said double stranded nucleotide sequence with the signal resulting from the double stranded nucleotide sequence formed by complementary base pairing between capture nucleotide sequence(s) and the target in order to quantify the presence of the original nucleotide sequence to be detected and/or quantified in the biological sample.

Advantageously the standard is added in the initial biological sample or after the extraction step and is amplified or copied with the same primers and/or has a length and a GC content identical or differing from no more than 20%

to the target. More preferably, the standard can be designed as a competitive internal standard having the characteristics of the internal standard found in the document WO98/11253. Said internal standard has a part of its sequence common to the target and a specific part which is different. It also has at or near its two ends sequences which are complementary of the two primers used for amplification or copy of the target and similar GC content (WO98/11253). In the preferred embodiment of this invention, the common part of the standard and the target, means a nucleotide sequence which is homologous to all target amplified by the same primers (i.e. which belong to the same family or organisms to be quantified).

Preferably, the hybridisation yield of the standard through this specific sequence is identical or differ no more than 20% from the hybridisation yield of the target sequence and quantification is obtained as described in WO 98/11253.

Said standard nucleotide sequence, external and/or internal standard, is also advantageously included in the kit (device) or apparatus according to the invention, possibly with all the media and means necessary for performing the different steps according to the invention (hybridisation and culture media, polymerase and other enzymes, standard sequence(s), labelling molecule(s), etc.).

Advantageously, the biochips also contain spots with various concentrations (i.e. 4) of labelled capture nucleotide sequences. These labelled capture nucleotide sequences are spotted from known concentrations solutions and their signals allow the conversion of the results of hybridisation into absolute amounts. They also allow to test for the reproducibility of the detection.

The solid support (biochip) can be inserted in a support connected to another chamber and automatic machine through the control of liquid solution based upon the use of

microfluidic technology. By being inserted into such a microlaboratory system, it can be incubated, heated, washed and labelled by automates, even for previous steps (like extraction of DNA, amplification by PCR) or the following step (labelling and detection). All these steps can be performed upon the same solid support.

The present invention will be described in details in the following non-limiting examples in reference to the enclosed figures.

Brief description of the drawings

Figure 1 is a schematic presentation of the step used in the method of the invention for the identification of 5 *Staphylococcus* species on biochips after PCR amplification with consensus primers.

Figure 2 represents the design of an array which allows the determination of the 5 most common *Staphylococcus* species, of the presence of any *Staphylococcus* strain and of the MecA gene.

Figure 3 presents the effect of the length of the specific sequence of a capture nucleotide sequence on the discrimination between sequences with different level of homology.

Figure 4 shows the sensitivity obtained for the detection of *FemA* sequences from S. *aureus* on array bearing the small specific capture nucleotide sequence for a *S. aureus* and a consensus sequence.

Example 1: Detection of homologous FemA sequences on array bearing long specific capture nucleotide sequences (Fig. 3)

Production of the capture nucleotide sequences and of the targets

The *FemA* genes corresponding to the different *Staphylococci* species were amplified separately by PCR using the following primers:

S. aureus 1 :	5' CTTTTGCTGATCGTGATGACAAA 3'
S. aureus 2 :	5' TTTATTTAAAATATCACGCTCTTCG 3'
S. epidermidis 1 :	5' TCGCGGTCCAGTAATAGATTATA 3'
S. epidermidis 2 :	5' TGCATTTCCAGTTATTTCTCCC 3'
S. haemolyticus 1 :	5' ATTGATCATGGTATTGATAGATAC 3'
S. haemolyticus 2 :	5' TTTAATCTTTTTGAGTGTCTTATAC 3'
S. saprophyticus 1 :	5' TAAAATGAAACAACTCGGTTATAAG 3'
S. saprophyticus 2 :	5' AAACTATCCATACCATTAAGTACG 3'
S. hominis 1 :	5' CGACCAGATAACAAAAAAGCACAA 3'
S. hominis 2 :	5' GTAATTCGTTACCATGTTCTAA 3'

The PCR was performed in a final volume of 50 µl containing: 1.5 mM $MgCl_2$, 10 mM Tris pH 8.4, 50 mM KCl, 0.8 µM of each primer, 50 µM of each dNTP, 50 µM of biotin-16-dUTP), 1.5 U of Taq DNA polymerase Biotools, 7.5% DMSO, 5 ng of plasmid containing *FemA* gene. Samples were first denatured at 94 °C for 3 min. Then 40 cycles of amplification were performed consisting of 30 sec at 94 °C, 30 sec at 60 °C and 30 sec at 72 °C and a final extension step of 10 min at 72 °C. Water controls were used as negative controls of the amplification. The sizes of the amplicons obtained using these primers were 108 bp for *S. saprophyticus*, 139 bp for *S. aureus*, 118 bp for *S. hominis*, 101 pb for *S. epidermidis* and 128 bp for *S. haemolyticus*. The sequences of the capture nucleotide sequences were the same as the corresponding amplicons but they were single strands.

The biochips also contains positive controls which were CMV amplicons hybridised on their corresponding capture nucleotide sequence and negative controls which were capture nucleotide sequences for a HIV-I sequence on which the CMV could not bind.

Capture nucleotide sequence immobilisation

The protocol described by Schena et al (Proc. Natl Acad. Sci. USA 93, 10614 (1996)) was followed for the grafting of aminated DNA to aldehyde derivatised glass. The aminated capture nucleotide sequences were spotted from solutions at concentrations ranging from 150 to 3000 nM. The capture nucleotide sequences were printed onto the silylated microscopic slides with a home made robotic device (250 μm pins from Genetix (UK) and silylated (aldehyde) microscope slides from Cell associates (Houston, USA)). The spots have 400 μm in diameter and the volume dispensed is about 0,5 nl. Slides were dried at room temperature and stored at 4 ^{o}C until used.

Hybridisation

At 65 μl of hybridisation solution (AAT, Namur, Belgium) were added 5 μl of amplicons and the solution was loaded on the array framed by an hybridisation chamber. For positive controls we added 2 nM biotinylated CMV amplicons of 437 bp to the solution; their corresponding capture nucleotide sequences were spotted on the array. The chamber was closed with a covership and slides were denatured at 95 ^{o}C for 5 min. The hybridisation was carried out at 60^{o} for 2 h. Samples were washed 4 times with a washing buffer.

Colorimetric detection

The glass samples were incubated 45 min at room temperature with 800 μl of streptavidin labelled with

colloidal gold 1000 x diluted in blocking buffer (Maleic buffer 100 mM pH 7.5, NaCl 150 mM, Gloria milk powder 0.1%). After 5 washes with washing buffer, the presence of gold served for catalysis of silver reduction using a staining revelation solution (AAT, Namur, Belgium). The slides were incubated 3 times 10 min with 800 µl of revelation mixture, then rinsed with water, dried and analysed using a microarray reader. Each slides were then quantified by a specific quantification software.

Fluorescence detection

The glass samples were incubated 45 min at room temperature with 800 µl of Cyanin 3 or Cyanin 5 labelled streptavidin. After washing the slides were dried before being stored at room temperature. The detection was performed in the array-scanner GSM 418 (Genetic Microsystem, Woburn, MA, USA) Each slide was then quantified by a specific quantification software.

The results give a cross-reaction between the species. For example, *epidermidis* amplicons hybridised on its capture probe give a value of 152, but give a value of 144, 9, 13 and 20 respectively for the S. *saprophyticus*, S. *aureus*, S. *haemolyticus* and S. *hominis* capture probes.

Example 2: Detection of homologous FemA sequences on array bearing small specific capture nucleotide sequences

Protocols for capture nucleotide sequences immobilisation and silver staining detection were described in example 1 but the capture nucleotide sequences specific of the 5 *Staphylococcus* species were spotted at concentrations of 600 nM and are the following :

Name	Sequence (5' -> 3')
Capture nucleotide sequence	
ATaur02	ATTTAAAATATCACGCTCTTCGTTTAG
ATepi02	ATTAAGCACATTTCTTTCATTATTTAG
AThae02	ATTTAAAGTTTCACGTTCATTTTGTAA
AThom02	ATTTAATGTCTGACGTTCTGCATGAAG
ATsap02	ACTTAATACTTCGCGTTCAGCCTTTAA

In this case, the targets are fragments of the FemA gene sequence corresponding to the different Staphylococci species which were amplified by a PCR using the following consensus primers :

APstap03: 5' CCCACTCGCTTATATAGAATTTGA 3'

APstap04: 5' CCACTAGCGTACATCAATTTTGA 3'

APstap05: 5' GGTTTAATAAAGTCACCAACATATT 3'

This PCR was performed in a final volume of 100 µl containing: 3 mM $MgCl_2$, 1 mM Tris pH 8, 1 µM of each primer, 200 µM of dACTP, dCTP and dGTP, 150 µM of dTTP, 50 µM of biotin-16-dUTP, 2,5 U of Taq DNA polymerase (Boehringer Mannheim, Allemagne), 1 U of Uracil-DNA-glycosylase heat labile (Boehringer Mannheim, Allemagne), 1 ng of plasmid containing *FemA* gene. Samples were first denatured at 94°C for 5 min. Then 40 cycles of amplification were performed consisting of 1 min at 94°C, 1 min at 50°C and 1 min at 72°C and a final extension step of 10 min at 72°C. Water controls were used as negative controls of the amplification. The sizes of the amplicons obtained using these primers were 489 bp for all species. Figure 4 shows only the results obtained with the amplicons for *S. epidermidis and S. xylosus*.

The hybridisation solution was prepared as in example 1 and loaded on the slides. Slides were denatured at 98°C for 5 min. Hybridisation are carried out at 50°C for 2h. Samples are then washed 4 times with a washing buffer. The values were very low and almost undetectable.

Example 3: Effect of the spacer length on the Sensitivity of detection of homologous FemA sequences on array bearing long capture nucleotide sequences with a small specific sequence

The experiment was conducted as described in example 2 with the same amplicons but the capture nucleotide sequences used are the following:

Name	Sequence (5' -> 3')
Capture nucleotide sequence Ataur02	ATTTAAAATATCACGCTCTTCGTTTAG
ATepi02	ATTAAGCACATTTCTTTCATTATTTAG
ATepi03	GAATTCAAAGTTGCTGAGAA ATTAAGCACATTTCTTTCATTATTTAG
ATepi04	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTAAGCACATTTCTTTCATTATTTAG
ATepi05	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCGTCT TCTTAAAATCTAAAGAA ATTAAGCACATTTCTTTCATTATTTAG

[a] The spacer sequences are underlined

The target amplicons were 489 bp long while the capture nucleotide sequences were 47, 67 or 87 bases single stranded DNA with a specific sequence of 27 bases.

Example 4: Specificity of the detection of FemA sequences from different bacterial species on the same array bearing long capture nucleotide sequences with a small specific sequence

The experiment was conducted as described in example 2 but the capture nucleotide sequences were spotted at concentrations of 3000 nM and are the following:

Name	Sequence (5' -> 3')
Capture nucleotide sequence	
Ataur27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTTAAAATATCACGCTCTTCGTTTAG
Atepi27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTAAGCACATTTCTTTCATTATTTAG
Athae27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTTAAAGTTTCACGTTCATTTTGTAA
Athom27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTTAATGTCTGACGTTCTGCATGAAG
Atsap27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ACTTAATACTTCGCGTTCAGCCTTTAA

a The spacer sequence is underlined. The specific sequences were of 27 bases

The targets are fragments of the FemA gene sequence corresponding to the different Staphylococci species which were amplified by PCR using the following consensus primers :

APcons3-1: 5' TAAYAAARTCACCAACATAYTC 3'

APcons3-2: 5' TYMGNTCATTTATGGAAGATAC 3'

A consensus sequence is present on the biochips which detects all the tested *Staphylococcus* species. All target sequences were amplified by PCR with the same pair of primers.

The size of the amplicons obtained using these primers were 587 bp for all species. The consensus sequence capture probe was a 489 base long single stranded DNA complementary to the amplicons of *S. hominis* as amplified in example 2. The detection was made in fluorescence. Homology between the consensus capture probe and the sequences of the femA from the 15 S. species were between 66 and 85%. All the sequences hybridized on this consensus capture probe.

Example 5: effect of the length of the specific sequence of the capture nucleotide sequence on the discrimination between homologous sequences (figure 3).

The experiment was conducted as described in example 4 but at a temperature of 43°C and the capture nucleotide sequences used are presented in the table here joined. The numbers after the names indicate the length of the specific sequences.

The *FemA* amplicons of *S. anaerobius* (a subspecies of *S. aureus*) were hybridised on an array bearing capture nucleotide sequences of 67 single stranded bases with either 15, 27 and 40 bases specific for the *S.aureus*, *anaerobius* and *epidermidis* at their extremities. The

difference between the capture nucleotide sequences of *anaerobius* and *aureus* was only one base in the 15 base capture nucleotide sequence and 2 in the 27 and the 40 bases.

	Sequence (5' -> 3')
Capture nucleotide sequence	
Ataur15	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCGTCTT CTTAAAATGCTCTTCGTTTAGTT
Ataur27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTTAAAATATCGCTCTTCGTTTAG
Ataur40	GAATTCAAAGTTGCTGAGAATAGTTCAAATCTTTATTTAAAATA TCACGCTCTTCGTTTAGTTCTTT
Atana15	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCGTCTT CTTAAAATGCTCTTCATTTAGTT
Atana27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCGGTTT AAAATATCACGCTCTTCATTTAG
Atana40	GAATTCAAAGTTGCTGAGAATAGTTCAAATCTTTGTTTAAAATA TCACGCTCTTCATTTAGTTCTTT
Atepi15	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCGTCTT CTTAAAATTTTCATTATTTAGTT
Atepi27	GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG ATTAAGCACATTTCTTTCATTATTTAG
Atepi40	GAATTCAAAGTTGCTGAGAATAGTTCAAATCTTTATTAAGCACA TTTCTTTCATTATTTAGTTCCTC

Example 6: Sensitivity of the detection of FemA sequences of Staphylococcus aureus on arrays bearing specific sequence as proposed by this invention and the consensus sequence (figure 4)

The experiment was conducted as described in example 4 with the capture nucleotide sequences spotted at concentrations of 3000 nM. The bacterial FemA sequences were serially diluted before the PCR and being incubated with the arrays.

Example 7: Detection of 16 homologous FemA sequences on array

The consensus primers and the amplicons were the same as described in the example 4 but the capture probes were chosen for the identification of 15 Staphylococcus species. The experiment is conducted as in example 4. The capture probes contain a spacer fixed on the support by its 5' end and of the following sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3' followed by the following specific sequences for the various femA from the different Stphylococcus.

S. aureus ATTTAAAATATCACGCTCTTCGTTTAG
S. epidermidis ATTAAGCACATTTCTTTCATTATTTAG
S. haemolyticus ATTTAAAGTTTCACGTTCATTTTGTAA
S. hominis ATTTAATGTCTGACGTTCTGCATGAAG
S. saprophyticus ACTTAATACTTCGCGTTCAGCCTTTAA

S. capitis ATTAAGAACATCTCTTTCATTATTAAG
S. caseolyticus ATAAAGACATTCGAGACGAAGGCT
S. cohnii ACTTAACACTTCACGCTCTGACTTGAG
S. gallinarum ACTTAAAACTTCACGTTCAGCAGTAAG
S. intermedius GTGGAAATCTTGCTCTTCAGATTTCAG

S. lugdunensis TTCTAAAGTTTGTCGTTCATTCGTTAG

S. schleiferi TTTAAAGTCTTGCGCTTCAGTGTTGAG

S. sciuri GTTGTATTGTTCATGTTCTTTTTCTAA

S. simulans TTCTAAATTCTTTTGTTCAGCGTTCAA

S. warneri AGTTAAGGTTTCTTTTTCATTATTGAG

S. xylosus GCTTAACACCTCACGTTGAGCTTGCAA

Example 8: Detection of 19 homogous *p34* Sequences of Mycobacteria

The *P34* genes present in all *Mycobacteria* are all amplified with the following consensus primers

Sens

MycU4 5' CATGCAGTGAATTAGAACGT 3' located at the position 496-515 of the gene, Tm = 56°C

Antisens

APmcon02 5' GTA**S**GTCAT**RRS**T**Y**CTCC 3' located at the position position 733-750 of the gene, Tm = 52-58°C

S = C or G

R = A or G

Y = T or C

The size of amplified products ranges from 123 to 258 bp

The following capture probes have been chosen for the specific capture of the Mycobacteria sequences.

Capture probes

Avium : 5' CGGTCGTCTCCGAAGCCCGCG 3' (21 nt)

Gastrii 1 : 5' GATCGGCAGCGGTGCCGGGG 3' (20 nt)

Gastrii 3 : 5' GTATCGCGGGCGGCAAGGT 3' (19 nt)
Gastrii 5 : 5' TCTGCCGATCGGCAGCGGTGCCGG 3' (24nt)
Gastrii 7 : 5' GCCGGGGCCGGTATTCGCGGGCGG 3' (24nt)
Gordonae : 5' GACGGGCACTAGTTGTCAGAGG 3' (22 nt)
Intracellulare 1: 5' GGGCCGCCGGGGGCCTCGCCG 3' (21 nt)
Intracellulare 3 : 5' GCCTCGCCGCCCAAGACAGTG 3' (21 nt)
Leprae: 5' GATTTCGGCGTCCATCGGTGGT 3' (22 nt)
Kansasi 1 : 5' GATCGTCGGCAGTGGTGACGG 3' (21 nt)
Kansasi 3 : 5' TCGTCGGCAGTGGTGAC 3' (17 nt)
Kansasi 5 : 5' ATCCGCCGATCGTCGGCAGTGGTGACG 3' (27 nt)
Malmoense : 5' GACCCACAACACTGGTCGGCG 3' (21 nt)
Marinum : 5' CGGAGGTGATGGCGCTGGTCG 3' (21 nt)
Scrofulaceum : 5' CGGCGGCACGGATCGGCGTC (20 nt)
Simiae: 5' ATCGCTCCTGGTCGCGCCTA 3' (20 nt)
Szulgai : 5' CCCGGCGCGACCAGCAGAACG 3' (21 nt)
Tuberculosis: 5' GCCGTCCAGTCGTTAATGTCGC 3' (22 nt)
Xenopi: 5' CGGTAGAAGCTGCGATGACACG 3' (22 nt)

Each of the sequences above comprises a spacer at its 5' end
Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'
Capture probes are aminated at their 5' end.

Example 9: Detection of *MAGE* genes

MAGE genes are all amplified with the following consensus primers

Sens

- DPSCONS2 5' GGGCTCCAGCAGCCAAGAAGAGGA 3', located at the 398-421 position of the gene
Tm = 78°C

Other amplicons have been added as sense primer in order to increase the efficiency of the PCR for some MAGEs

- DPSMAGE1 5' GGGTTCCAGCAGCCGTGAAGAGGA 3' Tm = 78°C

- DPSMAG8 5' GGGTTCCAGCAGCAATGAAGAGGA 3' Tm = 74°C
- DPSMAG12 5' GGGCTCCAGCAACGAAGAACAGGA 3' Tm = 76°C

Antisense

- DPASCONB4 5' CGGTACTCCAGGTAGTTTTCCTGC 3', located at the position 913-936 of the gene, Tm = 74°C

The size of the amplified products is around 530 bp

The following capture probes of 27 nucleotides have been chosen for the specific cpature of the MAGE sequences.

Capture probes

Mage 1 DTAS01 5' ACAAGGACTCCAGGATACAAGAGGTGC 3' Mage 2 DTAS02 5' ACTCGGACTCCAGGTCGGGAAACATTC 3'

Mage 3 DTS0306 5' AAGACAGTATCTTGGGGGATCCCAAGA 3'

Mage 4 DTAS04 5' TCGGAACAAGGACTCTGCGTCAGGCGA 3'

Mage 5 DTAS05 5' GCTCGGAACACAGACTCTGGGTCAGGG 3'

Mage 6 DTS06 5' CAAGACAGGCTTCCTGATAATCATCCT 3'

Mage 7 DTAS07 5' AGGACGCCAGGTGAGCGGGGTGTGTCT 3'

Mage 8 DTAS08 5' GGGACTCCAGGTGAGCTGGGTCCGGGG 3'

Mage 9 DTAS09 5' TGAACTCCAGCTGAGCTGGGTCGACCG 3'

Mage 10 DTAS10 5' TGGGTAAAGACTCACTGTCTGGCAGGA 3'

Mage 11 DTAS11 5' GAAAAGGACTCAGGGTCTATCAGGTCA 3'

Mage 12 DTAS12 5' GTGCTACTTGGAAGCTCGTCTCCAGGT 3'

Each of the sequences above comprises a spacer aminated at its 5' end in order to be covalently linked to the glass

Spacer sequence

5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'

They are spotted on aldehyde bearing glasses and used for the detection of the MAGEs amplified by the consensus primers given here above. The results show a non equivocal identification of the MAGEs present in the tumors

compared to identification using 12 specific PCR, one for each MAGE sequences.

Example 10: Identification of G-protein dopamin receptors subtypes in rat

Dopamine Receptor coupled to the G-protein are all amplified with the following consensus primers

Sens

- CONSENSUS2-3-4

5' TGCAGAC**M**ACCACCAACTACTT 3' located at the position 221-242 of the gene, Tm = 66°C

M = A or C

- CONSENSUS1-5

5' TG**M**GG**K**CCAAGATGACCAAC**WT** 3' (22 nt) located at the position 221-240 of the gene, Tm = 66°C

M = A or C

K = G or T

W = A or T

Antisens

5 TCATG**R**C**R**CA**S**AGGTTCAGGAT 3' located at the position 395-416 of the gene, Tm = 64-68°C

R = A or G

S = C or G

The size of the amplified product is 196 pb.

The following capture probes of 27 nucleotides have been chosen for the specific capture of the dopamine receptor sequences.

Capture probes

DRD1 5' CTGGCTTTTGGCCTTTGGGTCCCTTTT 3' DRD2 5'

TGATTGGAAATTCAGCAGGATTCACTG 3' DRD3 5'

GAGTCTGGAATTTCAGCCGCATTTGCT 3' DRD4 5'

CGTCTGGCTGCTGAGCCCCCGCCTCTG 3' DRD5 5'

CTGGGTACTGGCCCTTTGGGACATTCT 3'

Each of the sequences above comprises an aminated spacer at its 5' end.Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG

Example 11: Identification of G-protein histamin receptors subtypes in rat

Histamin Receptor coupled to the G-protein are all amplified with the following primers

Sens

- H1sens

5' CTCCGTCCAGCAACCCCT 3' (18 nt) located at the Position 381-398 of the gene, Tm = 60°C

- H2sens

5' CTGTGCTGGTCACCCCAGT 3' (18 nt) located at the Position 380-398 of the gene, Tm = 62°C

- H3sens

5' ACTCATCAGCTATGACCGATT 3'(21 nt) located at the Position 378-398 of the gene, Tm = 60°C

Antisens

- H1antisens

5' ACCTTCCTTGGTATCGTCTG 3'(20 nt) located at the Position 722-741 of the gene, Tm = 60°C

- H2antisens

5' GAAACCAGCAGATGATGAACG 3'(21 nt) located at the Position 722-742 of the gene, Tm = 62°C

- H3antisens

5' GCATCTGGTGGGGGTTCTG 3'(19 nt) located at the Position 722-740 of the gene, Tm = 62°C

Size of the amplified product ranges from 359 to 364 pb.

The following capture probes have been chosen for the specific capture of the histamin receptor sequences.

Capture probes

H1 5' CCCCAGGATGGTAGCGGA 3' (18 nt)

H2 5' AGGATAGGGTGATAGAAATAAC 3' (22 nt)

H3 5' TCTCGTGTCCCCCTGCTG 3' (18 nt)

Each of the sequences above comprises a spacer at its 5' end

Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'. Capture probes are aminated at their 5' end.

Example 12: Identification of G-protein serotonin receptors subtypes in rat

Serotonin Receptor coupled to the G-protein are all amplified with the following primers

Sens

- Consensus for the subtypes 1A, 1B, 1C, 1D, 1E, 2A, 2B, 2C, 4, 6, 7

5'ATC**H**TGCACCT**S**TG**B**G**B**CAT 3' Tm = 58-64°C (20 nt)

H = C or A or T

S = C or G

B = C or T or G

1A ATCCTGCACCTGTGCGCCAT (0 mismatch) position 370-389

1B ATCATGCATCTCTGTGTCAT (1 mismatch) position 397-416

1C ATCATGCACCTCTGCGCCAT (0 mismatch) position 427-446

1D ATCCTGCATCTCTGTGTCAT (1 mismatch) position 367-386

1E ATCTTGCACCTGTCGGCTAT (2 mismatch) position 331-350

2A ATCATGCACCTCTGCGCCAT (0 mismatch) position 487-506

2B ATCATGCATCTCTGTGCCAT (1 mismatch) position 424-443

2C ATCATGCACCTCTGCGCCAT (0 mismatch) position 24-43

4 ATTTTTCACCTCTGCTGCAT (3 mismatchs)

6 ATCCTCAACCTCTGCTTCAT (3 mismatchs)

7 ATCATGACCCTGTGCGTGAT (3 mismatchs)

- Consensus 4, 6

5' ATCYTYCACCTCTGCYKCAT 3' Tm = 52-64°C (20 nt)

K = G or T

Y = T or C

4 ATTTTTCACCTCTGCTGCAT (1 mismatch) position 322-341

6 ATCCTCAACCTCTGCCTCAT (1 mismatch) position 340-359

- Consensus 5A, 5B

5' ATCTGGAAYGTGRCAGCCAT 3' Tm = 58-62°C (20 nt)

Y = T or C

R = A or G

5A ATCTGGAATGTGACAGCAAT (1 mismatch) position 385-404

5B ATCTGGAACGTGGCGGCCAT (1 mismatch) position 424-443

- Spécifique 7

5' ATCATGACCCTGTGCGTGAT 3' Tm = 56°C (18 nt) position 517-536

- Spécifique 3B

5' CTTCCGGAACGATTAGAAA 3' Tm = 54°C (19 nt) position 404-422

Antisens

- Consensus for the subtypes 1A, 1B, 1C, 1D, 1E, 2A, 2B, 2C, 4, 7 Tm = 48-58 °C

5' TTGG**HN**GC**Y**TTC**Y**G**B**TC 3'

H = A or T or C

N = A or C or G or T

B = C or T or G

1A TTCACCGTCTTCCTTTC (4 mismatchs)

1B TTGGTGGCTTTGCGCTC (1 mismatch) position 913-929

1C TTGGAAGCTTTCTTTTC (1 mismatch) position 922-938

1D TTAGTGGCTTTCCTTTC (2 mismatchs) position 877-893

1E GTGGCTGCTTTGCGTTC (2 mismatchs) position 862-878

2A TTGCACGCCTTTTGCTC (2 mismatchs) position 952-968

2B TTTGAGGCTCTCTGTTC (2 mismatchs) position 952-968

2C TTGGAAGCTTTCTTTTC (1 mismatch) position 424-440

4 TTGGCTGCTTTCCGGTC (2 mismatchs)

7 GTGGCTGCTTTCTGTTC (1 mismatch) position 973-989

- Specific 1A

5' TTCACCGTCTTCCTTTC 3' Tm = 50°C (17 nt) position 1018-1034

- Specific 4

5' TCTTGGCTGCTTTGGTC 3' Tm = 52°C (17 nt) position 762-778

- Specific 6

5' ATAAAGAGCGGGTAGATG 3' Tm = 52°C (18 nt) position 945-963

- Consensus 5A, 5B

5' CCTTCTGCTCCCTCCA 3' Tm = 52°C (16 nt)

5A CCTTCTGTTCCCTCCA (1 mismatch) position 823-840

5B CCTTCTGCTCCCGCCA (1 mismatch) position 862-879

- Specific 3B

5' ACCGGGGACTCTGTGT 3' Tm = 52°C (16 nt) position 1072-1089

The following capture probes have been chosen for the specific capture of the serotonin receptor subtypes sequences.

Capture probes

HTR1C 5' CTATGCTCAATAGGATTACGT 3' (21 nt)
HTR2A 5' GTGGTGAATGGGGTTCTGG 3' (19 nt)
HTR2B 5' TGGCCTGAATTGGCTTTTTGA 3' (21 nt)
HTR2C/1C 5' TTATTCACGAACACTTTGCTTT 3' (22 nt)
HTR1B 5' AATAGTCCACCGCATCAGTG 3' (20 nt)
HTR1D 5' GTACTCCAGGGCATCGGTG 3' (19 nt)
HTR1A 5' CATAGTCTATAGGGTCGGTG 3' (20 nt)
HTR1E 5' ATACTCGACTGCGTCTGTGA 3' (20 nt)
HTR7 5' GTACGTGAGGGGTCTCGTG 3' (19 nt)
HTR5A 5' GGCGCGTTATTGACCAGTA 3' (19 nt)
HTR5B 5' GGCGCGTGATAGTCCAGT 3' (18 nt)
HTR3B 5' GATATCAAAGGGGAAAGCGTA 3' (21 nt)
HTR4 5' AAACCAAAGGTTGACAGCAG 3' (20 nt)
HTR6 5' GTAGCGCAGCGGCGAGAG 3' (18 nt)

Each of the sequences above comprises a spacer at its 5' end

Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'. Capture probes are aminated at their 5' end.

Example 13 : Identification of the HLA-A subtypes

The HLA-A subtypes are amplified with the following consensus primers

Sens

IPSCONA 5' GACAGCGACGCCGCGAGCCA 3' located at the position 181-200 of the gene, Tm = 70°C

Antisens

IPASCONA 5 CGTGTCCTGGGTCTGGTCCTCC 3' located at the position 735-754 of the gene, Tm = 74°C

The size of the amplified product is 574 bp

The following capture probes of 27 nucleotides have been chosen for the specific capture of the HLA-A sequences

Capture probes

HLA-A1	ITSA01	5' GGAGGGCCGGTGCGTGGACGGGCTCCG 3'
HLA-A2	ITASA02	5' TCTCCCCGTCCCAATACTCCGGACCCT 3'
HLA-A3	ITASA03A	5' CTGGGCCTTCACATTCCGTGTCTCCTG 3'
	ITSA03B	5' AGCGCAAGTGGGAGGCGGCCCATGAGG 3'
HLA-A11	ITSA11A	5' GCCCATGCGGCGGAGCAGCAGAGAGCC 3'
	ITSA11B	5' CCTGGAGGGCCGGTGCGTGGAGTGGCT 3'
HLA-A23	ITSA23A	5' GCCCGTGTGGCGGAGCAGTTGAGAGCC 3'
	ITASA23B	5' CCTTCACTTTCCCTGTCTCCTCGTCCC 3'
HLA-A24	ITSA24A	5' GCCCATGTGGCGGAGCAGCAGAGAGCC 3'
	ITASA24B	5' TAGCGGAGCGCGATCCGCAGGTTCTCT 3'
HLA-A25	ITASA25A	5' TAGCGGAGCGCGATCCGCAGGCTCTCT 3'
	ITASA25B	5' TCACATTCCGTGTGTTCCGGTCCCAAT 3'
HLA-A26	ITASA26	5' GGGTCCCCAGGTTCGCTCGGTCAGTCT 3'
HLA-A29	ITASA29	5' TCACATTCCGTGTCTGCAGGTCCCAAT 3'
HLA-A30	ITASA30	5'CGTAGGCGTGCTGTTCATACCCGCGGA 3'
HLA-A31	ITASA31	5' CCCAATACTCAGGCCTCTCCTGCTCTA
3'HLA-A33	ITSA33	5' CGCACGGACCCCCCCAGGACGCATATG 3'
HLA-A68	**ITSA68A**	**5' GGCGGCCCATGTGGCGGAGCAGTGGAG 3'**
	ITASA68B	5' GTCGTAGGCGTCCTGCCGGTACCCGCG 3'
HLA-A69	ITASA69	5' ATCCTCTGGACGGTGTGAGAACCGGCC 3'

Each of the sequences above comprises an aminated spacer at its 5' end.Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'

Example 14: Identification of Cytochrome P450 3a forms

The Cytochrome P450 forms are amplified with the followingconsensus primers

Sens

- Consensus

5' GCCAGAGCCTGAGGA 3' located at the position 1297-1311 of the 3a3 gene, Tm = 50°C

Antisens

- Consensus a3, a23, a1, a2

5' TCAAAAGAAATTAACAGAGA 3' located at the position 1839-1858 of the 3a3 gene, Tm = 50°C

- Specific a9

5' ACAATGAAGGTAACATAGG 3' located at the position 2015-2033 of the 3a9 gene Tm = 52°C

- Specific a18

5' ACTGATGGAACTAACTGG 3' located at the position 1830-1846 of the 3a18 gene Tm = 52°C

The length of the PCR product is around 560pb.

The following capture probes have been chosen for the specific capture of the cytochrome P-450 3a sequences.

Capture probe

3a1 5' TGTTTTGATTCGGTACATCTTTG 3' (24 nt)

3a3 5' TTGATTTGGTACATCTTTGCT 3' (21 nt)

3A9 5' ACTCCTGGGGGTTTTGGGTG 3' (20 nt)

3A18 5' ATTACTGAGTATTCAGAAATTCAC 3' (24 nt)

3A2 5' GGTTAAAGATTTGGTACATTTATGG 3' (25 nt)

Each of the sequences above comprises a spacer at its 5' end

Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG 3'. Capture probes are aminated at their 5' end.

Example 15 : Identification of GMO on biochips

Consensus primers to detect GMO on biochips:

OGM1 CGTCTTCAAAGCAAGTGGATTG

OGM2 ATCCTGTTGCCGGTCTTGCG

These primers allow the amplification of the genes:

1) CTP1, CTP2, CP4EPSPS, S CryIAb and hsp 70 Int. in Mon 809 (corn, Monsanto)

2) hsp 70 Int. and S CryIAb in Mon 810 (corn, Monsanto)

3) S CryIAb and S Pat in Bt 11 (corn, Novartis)

4) CTP4 and EPSPS in GTS40-3-2 (soybean, Monsanto)

The capture probes will be chosen in these sequences to allow discrimination. Each of the sequences above comprises a spacer at its 5' end

Spacer sequence 5' GAATTCAAAGTTGCTGAGAATAGTTCAATGGAAGGAAGCG